SYMBOLIC SOLUTION OF CARD MATCHING PROBLEMS

N. S. MENDELSOHN

The main problem to be discussed here is the following. Find the number of arrangements of n cards marked $1, 2, \cdots, n$ subject to conditions of the type: the card marked "i" shall not be jth, the card marked "k" shall not be rth, and so on. A generalization of this problem is also discussed.

A solution of the card matching problem has been given by Kaplansky in [2].[1] The present solution depends on a somewhat different approach to the problem. Both Kaplansky and I make use of the finite difference operator E, defined by $Ef(n)=f(n+1)$: Kaplansky's solution is based on a symbolic interpretation of the method of inclusion and exclusion; my solution gives a recurrence formula expressing the solution of the problem of matching n cards in terms of the solution of the problems of matching less than n cards. The solution proposed here is capable of giving explicit formulae for several particular cases, for example, the "problème des ménages." Furthermore, it is capable of being extended to problems of considerably greater generality.

Suppose we have $a_1, a_2, \cdots, a_n$ cards, all considered distinct, of which a_r are marked r. It is required to find the number of arrangements of these cards in which none of the cards marked "r" appear in any of p_r specified places. As an immediate corollary, we also obtain the number of arrangements in which these conditions are violated (1) exactly s times and (2) at most s times.

Let p_{rs} be the number of places simultaneously forbidden to cards marked r or s; p_{rst} the number of places simultaneously forbidden to cards marked r, s or t, and so on. The form our solution takes depends on the $p_{rst} \ldots$ with the largest number of subscripts which does not vanish. We give the following examples.

Case I. All $p_i=0$. The number of suitable arrangements is $E^{a_1+\cdots+a_n}0!$. This is obvious.

Case II. Some $p_i\neq 0$, but all $p_{ij}=0$. The number of suitable arrangements is $F_1(a_1;\ p_1)F_1(a_2;\ p_2)\cdots F_1(a_n;\ p_n)0!$ where $F_1(a;\ p) =\sum(-1)^r[a, r][p, r]E^{a-r}$, the summation being carried out with respect to r which ranges from 0 to min (a, p). The symbol $[a, r]$ is used

Received by the editors July 17, 1945, and, in revised form, December 17, 1945, and January 4, 1946.

[1] Numbers in brackets refer to the references cited at the end of the paper.

for the binomial coefficient $a(a-1)\cdots(a-r+1)/r!$. This problem was solved by Kaplansky in [1].

Case III. Some $p_{ij}\neq 0$, all $p_{ijk}=0$. The number of suitable arrangement is $F_n(a_1,\cdots,a_n;p_1,\cdots,p_n;p_{12},\cdots,p_{1n},p_{23},\cdots,p_{n-1,n})0!$ (written as $F_n(a_i;p_i;p_{ij})0!$ if no ambiguity occurs). Here $F_1(a;p)$ is the same as defined in case II while $F_n(a_i;p_i;p_{ij})$ is defined inductively by the recurrence

$$F_n(a_i;p_i;p_{ij}) = \sum(-1)^{k_1+\cdots+k_{n-1}}k_1!\cdots k_{n-1}![a_1,k_1][p_{1n},k_1]\cdots[a_{n-1},k_{n-1}][p_{n-1,n},k_{n-1}]BC$$

where the summation is carried out with respect to each of the indices $k_1, k_2, \cdots, k_{n-1}$, the index k_r ranging from 0 to min (a_r, p_{rn}). Here B and C represent the following expressions: $B=F_{n-1}(a_r-k_r;p_r-p_{rn};p_{rs})$ where r and s range from 1 to $n-1$; $C=F_1(a_n;p_n-\{k_1+\cdots+k_{n-1}\})$.

Case IV. Some $p_{ijk}\neq 0$, all $p_{ijkl}=0$. The number of suitable arrangements is $F_n(a_i;p_i;p_{ij};p_{ijk})0!$ $(i, j, k=1, 2, \cdots, n; i<j<k)$ where $F_n(a_i;p_i;p_{ij};p_{ijk})=\sum(-1)^{k_1+\cdots+k_{n-2}}k_1!\cdots k_{n-2}![a_1,k_1]\cdot[p_{1,n-1,n},k_1]\cdots[a_{n-2},k_{n-2}][p_{n-2,n-1,n},k_{n-2}]F_n^*$, where the summation is carried out with respect to each of the indices $k_1, k_2, \cdots, k_{n-2}$. Here k_r ranges from 0 to min $(a_r, p_{r,n-1,n})$ and $F_n^*=F_n(a_i^*;p_i^*;p_{ij}^*;p_{ijk}^*)$, where:

$$a_i^* = a_i - k_i \qquad (i=1,2,\cdots,n-2),$$

$$a_{n-1}^* = a_{n-1}, \qquad a_n^* = a_n,$$

$$p_i^* = p_i - p_{i,n-1,n} \qquad (i=1,2,\cdots,n-2),$$

$$p_{n-1}^* = p_{n-1} - (k_1+k_2+\cdots+k_{n-2}),$$

$$p_n^* = p_n - (k_1+k_2+\cdots+k_{n-2}),$$

$$p_{ij}^* = p_{ij} \qquad (i,j=1,2,\cdots,n-2; j>i),$$

$$p_{i,n-1}^* = p_{i,n-1} - p_{i,n-1,n} \qquad (i=1,2,\cdots,n-2),$$

$$p_{in}^* = p_{in} - p_{i,n-1,n} \qquad (i=1,2,\cdots,n-2),$$

$$p_{n-1,n}^* = p_{n-1,n} - (k_1+k_2+\cdots+k_{n-2}),$$

$$p^*_{ijk} = p_{ijk} \qquad \text{unless } j = n-1,\ k = n,$$

$$p^*_{i,n-1,n} = 0.$$

Note that this gives a reduction formula in which $p_{i,n-1,n}$ are reduced to 0. By permuting the a_i, repeated applications of the above formula reduce all the p_{ijk} to 0, after which the formulas of case III apply.

The general case can now be written down by analogy.

If we wish to find the number of arrangements in which the conditions are violated (1) exactly s times and (2) at most s times, we replace $0!$ by $(-1)^s[m-0,\ s]0!$ and by $(-1)^s[m-1-0,\ s]0!$ respectively, where $m=\sum a_i$ (see [2] and [4]).

We give a proof for the case III; the extension of the proof to the other cases is obvious.

In the formulas as they stand certain conventions must be made for the cases where some a_i or p_i reduce to 0. These are:

$$F_n(0, 0, \cdots, 0; p_1, \cdots, p_n; p_{12}, \cdots, p_{n-1,n}) = 1,$$

$$F_n(0, a_2, \cdots, a_n; p_1, \cdots, p_n; p_{12}, \cdots, p_{n-1,n}) = F_{n-1}(a_2, \cdots, a_n; p_2, \cdots, p_n; p_{23}, \cdots, p_{n-1,n}),$$

$$F_n(a_1, \cdots, a_n; 0, p_2, \cdots, p_n; p_{12}, \cdots, p_{n-1,n}) = E^{a_1}F_{n-1}(a_2, \cdots, a_n; p_2, \cdots, p_n; p_{23}, \cdots, p_{n-1,n}).$$

With these conventions the proof of the theorem is obtained by induction on the sum $\sum p_{ij}$. If $\sum p_{ij}=0$, the formula quoted for F_n in case III becomes

$$F_n(a_1, \cdots, a_n; p_1, \cdots, p_n; 0, \cdots, 0) = F_{n-1}(a_1, \cdots, a_{n-1}; p_1, \cdots, p_{n-1}; 0, \cdots, 0)F_1(a_n; p_n).$$

Hence by induction,

$$F_n(a_1, \cdots, a_n; p_1, \cdots, p_n; 0, \cdots, 0) = \prod F_1(a_i; p_i),$$

the product ranging over all terms $i=1, 2, \cdots, n$.

This is precisely the formula quoted in case II and was given by Kaplansky in [1]. We write the number of suitable arrangements in the form

$$F_n(a_1, \cdots, a_n; p_1, \cdots, p_n; p_{12}, \cdots, p_{n-1,n})0!.$$

Suppose now that $p_{n-1,n}\neq 0$. (This can be brought about by a change of notation if necessary, since we are now considering the case where not all $p_{ij}=0$.) Consider one of the places which is forbidden simultaneously to cards marked $(n-1)$ and to cards marked n. If we lift the prohibition on the cards marked $(n-1)$ to this place, the number of suitable arrangements is

$$F_n(a_1, \cdots, a_{n-1}, a_n; p_1, \cdots, p_{n-1}-1, p_n; p_{12}, \cdots, p_{n-1,n}-1)0!.$$

Of these arrangements,

$$a_{n-1}F_n(a_1, \cdots, a_{n-1}-1, a_n; p_1, \cdots, p_{n-1}-1, p_n-1; \\ p_{12}, \cdots, p_{n-1,n}-1)0!$$

have a card marked $(n-1)$ in the specified place. Thus

$$\begin{aligned} &F_n(a_1, \cdots, a_n; p_1, \cdots, p_n; p_{12}, \cdots, p_{n-1,n}) \\ &\quad = F_n(a_1, \cdots, a_n; p_1, \cdots, p_{n-1}-1, p_n; p_{12}, \cdots, p_{n-1,n}-1) \\ &\qquad - a_{n-1}F_n(a_1, \cdots, a_{n-1}-1, a_n; p_1, \cdots, p_{n-1}-1, p_n-1; \\ &\qquad\qquad p_{12}, \cdots, p_{n-1,n}-1). \end{aligned}$$

The induction is established by substituting the values in the right-hand side of this equation into the recurrence stated for F_n.

For computing purposes, the following remark is useful. If $p_{ij}=0$ $(i=1, 2, \cdots, r; j=r+1, r+2, \cdots, n)$, then

$$\begin{aligned} &F_n(a_1, \cdots, a_r, a_{r+1}, \cdots, a_n; p_1, \cdots, p_r, p_{r+1}, \cdots, p_n; p_{12}, \cdots, p_{n-1,n}) \\ &\quad = F_r(a_1, \cdots, a_r; p_1, \cdots, p_r; p_{12}, \cdots, p_{r-1,r}) \\ &\qquad \cdot F_{n-r}(a_{r+1}, \cdots, a_n; p_{r+1}, \cdots, p_n; p_{r+1,r+2}, \cdots, p_{n-1,n}). \end{aligned}$$

This factorization holds in the most general case and is obvious from first principles, or it can be established by induction.

We now give an example illustrating the use of the above formulae. In this example, the condition

$$\sum a_i \geqq \sum p_i - \sum p_{ij} + \sum p_{ijk} \cdots,$$

which is necessary if the problem as stated is to have a meaning, is not satisfied. The result obtained is still useful for two reasons: (1) the problem may be made meaningful by embedding it in a larger problem, for example, by adding $\sum p_i - \sum p_{ij} + \sum p_{ijk} - \cdots - \sum a_{ij}$ blank cards with no restriction as to where these cards are to be put; (2) the coefficient of E^{n-k} in the polynomial $f(E)$ gives the number of ways of choosing k compatible conditions from the converse conditions and this is always meaningful. (The converse condition to "i is not jth" is defined as the condition "i is jth.")

Example. Let $a_i=1$; $p_i=p$ $(i=1, 2, \cdots, n)$; $p_{12}=p_{23}= \cdots =p_{n-1,n} =r$; all remaining $p_{ij}=0$; all $p_{ijk}=0$. These conditions imply $r \leqq p/2$. Let the polynomial operator be U_n. By applying the recurrence formula of case III twice one obtains

$$U_n = (E - p)U_{n-1} - {}_rU_{n-2}.$$

Also, $U_1 = E - p$; $U_2 = E^2 - 2pE + (p^2 - r)$. By induction it is easily shown that

$$U_n = \sum_{k=0}^{n} (-1)^k \{ [n, 0][n, k]p^k + \cdots + (-1)^s[n - s, s][n - 2s, k - 2s]p^{k-2s}r^s + \cdots \} E^{n-k}.$$

In the particular case where $p = 2$, $r = 1$, the solution can be reduced to the simpler form

$$U_n = \sum_{k=0}^{n} (-1)^k [2n - k - 1, k] E^{n-k}.$$

This polynomial is useful in the solution of the "problème des ménages"; cf. [3].

Generalization to the matching of several decks of cards. Battin [5] and Wilks [6] have discussed the following problem. Suppose there are k decks of cards of several suits $C_1, C_2, \cdots, C_r$, the decks not necessarily having the same number of cards in each suit, nor need the decks be of identical composition. One card at a time is dealt from each deck, and if all the cards at a drawing are of the same suit a k-ple match is said to occur. Amongst the questions of interest are: (1) the probability of h *k-ple* matches (simply called matches from here on), and (2) the probability of at most h matches.

A case which is very easily treated by symbolic methods is that of $(s+1)$ decks of cards, s of which are identical, the other deck being different. We also assume that this odd deck has less cards than the other decks. Let the identical decks have a_1 cards of suit 1, a_2 cards of suit 2, $\cdots$, a_n cards of suit n and let $a_1 + a_2 + \cdots + a_n = N$. Let the remaining deck have p_i cards of suit i ($i = 1, 2, \cdots, n$), where $\sum p_i \leqq \sum a_i = N$. We can show that the probability of no matches is given by

$$P(0) = \frac{1}{(N!)^s} \left\{ \prod_{i=1}^{n} \Phi_1^{(s)}(a_i; p_i) \right\} (0!)^s,$$

the probability of exactly h matches is given by

$$P(h) = \frac{1}{(N!)^s} \left\{ \prod_{i=1}^{n} \Phi_1^{(s)}(a_i; p_i) \right\} (-1)^h [N - 0, h](0!)^s,$$

while the probability $Q(h)$ of at most h matches is given by

$$Q(h) = \frac{1}{(N!)^s}\left\{\prod_{i=1}^{n} \Phi_1^{(s)}(a_i; p_i)\right\}(-1)^h[N-0-1, h](0!)^s,$$

where $\Phi_1^{(s)}(a; p) = \sum(-1)^i(i!)^s[a, i]^s[p, i]E^{a-i}$, the summation being with respect to i which ranges from 0 to min (a, p). The proof of this can be obtained by induction on observing that $\Phi_1^{(s)}(a; p)$ satisfies the recurrence

$$\Phi_1^{(s)}(a; p) = \Phi_1^{(s)}(a; p-1) - a^s\Phi_1^{(s)}(a-1, p-1).$$

Note that the case $s=1$ (that is, the general two deck case) has first been given by Kaplansky and Fréchet [1 and 4], although a solution by Greville [7] is usually quoted as the first obtained.

We can even generalize the card matching problem as follows.

Suppose that there are s identical packs of cards, each pack having a_i cards, considered as distinct, marked i $(i=1, 2, \cdots, n)$. Suppose that the packs be arranged in a rectangular array, each row of the array consisting of all the cards of one pack. Suppose that to cards marked i, we prohibit p_i columns in the sense that a column forbidden to cards marked i must not contain s cards marked i (although less than s cards marked i is permissible). It is required to find the number of suitable arrangements.

Let p_{ij} be the number of columns forbidden simultaneously to cards marked i and j. Similarly p_{ijk}, p_{ijkl}, and so on, are defined. We consider only the analogue of case III of our original problem, extension to other cases being obvious. Let $p_{ijk}=0$ for all i, j, k, but at least one $p_{ij}\neq 0$. The number of suitable arrangements is

$$\Phi_n^{(s)}(a_1, \cdots, a_n; p_1, \cdots, p_n; p_{12}, \cdots, p_{n-1,n})(0!)^s,$$

where $\Phi_n^{(s)}$ satisfies the following recurrence: $\Phi_1^{(s)}(a, p)$ is identical with previously given definition;

$$\begin{aligned}\Phi_n^{(s)}&(a_1, \cdots, a_n; p_1, \cdots, p_n; p_{12}, \cdots, p_{n-1,n})\\ &= \sum(-1)^{k_1+\cdots+k_{n-1}}(k_1!\cdots k_{n-1}!)^s[a_1, k_1]^s[p_{1n}, k_1]\\ &\qquad\cdots[a_{n-1}, k_{n-1}]^s[p_{n-1,n}, k_{n-1}]BC,\end{aligned}$$

where the summation is carried out with respect to each of the indices $k_1, k_2, \cdots, k_{n-1}$, the index k_r ranging from 0 to min (a_r, p_{rn}). Here B and C represent the following expressions:

$$B = \Phi_{n-1}^{(s)}(a_r - k_r; p_r - p_{rn}; p_{rs}),$$

where r and s range from 1 to $n-1$; $C = \Phi_1^{(s)}(a_n; p_n - \{k_1+k_2+\cdots+k_{n-1}\})$.

The proof of this formula is by induction on $\sum p_{ij}$, where we make use of the relationship $\Phi_n^{(s)}(a_1, \cdots, a_n; p_1, \cdots, p_n; p_{12}, \cdots, p_{n-1,n}) = \Phi_n^{(s)}(a_1, \cdots, a_n; p_1, \cdots, p_{n-1}-1, p_n; p_{12}, \cdots, p_{n-1,n}-1) - a_{n-1}^s\Phi_n^{(s)}(a_1, \cdots, a_{n-1}-1, a_n; p_1, \cdots, p_{n-1}-1, p_n-1; p_{12}, \cdots, p_{n-1,n}-1)$.

The probability of (1) exactly h violations of the conditions and (2) at most h violations of the conditions can be written down in the usual way. If all the $a_i=1$, then $\Phi_n^{(s)}=F_n$. It follows that if $f(E)0!$ is a solution of a one-pack problem then $f(E)(0!)^s$ is the solution of an s-pack problem provided $a_i=1$.

For example, the problème des ménages asks for the number of ways in which n cards marked 1, 2, $\cdots$, n, respectively, can be arranged so that the card marked 1 is not first or second, the card marked 2 is not second or third . . . the card marked n is not nth or first. The solution is $M_n(E)0!$ where

$$M_n(E) = \sum_{k=0}^{n} (-1)^k \left(\frac{2n}{2n-k}\right) [2n-k, k]E^{n-k}.$$

Since all the $a_i=1$, the solution to a s-pack problème des ménages, in the sense of our theorem, is $M_n(E)(0!)^s$.

An alternative approach to the multiple matching problems, using symbolic methods, has recently been given by Kaplansky and Riordan in [8].

References

1. I. Kaplansky, *Symbolic solution of certain problems in permutations*, Bull. Amer. Math. Soc. vol. 50 (1944) pp. 906–914.

2. ———, *On a generalization of the "problème des rencontres,"* Amer. Math. Monthly vol. 46 (1939) pp. 159–161.

3. ———, *Solution of the "problème des menâges,"* Bull. Amer. Math. Soc. vol. 49 (1943) pp. 784–785.

4. M. Fréchet, *Note on the "problème des rencontres,"* Amer. Math. Monthly vol. 46 (1939) p. 501.

5. I. L. Battin, *On the problem of multiple matching*, Ann. Math. Statist. vol. 13 (1942) pp. 294–305.

6. S. S. Wilks, *Mathematical statistics*, pp. 208–213.

7. T. N. E. Greville, *The frequency distribution of a general matching problem*, Ann. Math. Statist. vol. 12 (1941) pp. 350–354.

8. I. Kaplansky and J. Riordon, *Multiple matching and runs by the symbolic method*, Ann. Math. Statist. vol. 16 (1945) pp. 272–277.

QUEEN'S UNIVERSITY, KINGSTON, ONT.